Meta-Analysis versus Large Clinical Trials: Which Should Guide Our Management?

Christina M. Scifres, M.D.[1], **Jay D. Iams, M.D.**[2], **Mark Klebanoff, M.D., M.P.H.**[3], and **George A. Macones, M.D., M.S.C.E.**[1]

[1]Department of Obstetrics and Gynecology Washington University in St. Louis

[2]Department of Obstetrics and Gynecology The Ohio State University

[3]Division of Epidemiology, Statistics and Prevention Research, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health

Abstract

Large, randomized clinical trials have long been considered the gold standard to guide clinical care. Meta-analysis is a type of analysis in which results of a number of randomized clinical trials are combined and a summary measure of effect for a given treatment is ascertained. The clinician in practice is often faced with a dilemma regarding the type of evidence that should be used to guide clinical practice and for many clinical problems there are both randomized controlled trials (RCTs) and meta-analyses available. The cases of calcium and aspirin for the prevention of pre-eclampsia afford an opportunity to explore the benefits and limitations of each type of study to guide clinical practice. We conclude that when available, large, randomized clinical trials should be used to guide clinical practice.

Keywords

Professional societies and government agencies are reluctant to promulgate or endorse guidelines for medical care until clinical trials produce a consensus. When consensus cannot be achieved because only small trials are available or trial outcomes vary, meta-analysis is often performed with the expectation that larger numbers will allow a consensus to emerge. In some situations, both large randomized controlled trials (RCTs) and meta-analyses are available. In this case, which should guide clinical care, large clinical trials, or meta-analysis? In this essay, we compare the strengths and weaknesses of each, using two

Corresponding Author: Christina M. Scifres, M.D., Washington University School of Medicine, Department of Obstetrics and Gynecology, 4566 Scott Avenue, Box 8064, St. Louis, MO 63110, Office Phone (314)362-7300, Cell Phone (314) 249-6612, Fax (314) 362-0041, scifresc@wudosis.wustl.edu.

Reprints not available

proposed therapies for the prevention of pre-eclampsia as current examples of this conundrum.

Pre-eclampsia is a common illness affecting pregnant women. It manifests clinically as new-onset elevated blood pressure and proteinuria in the second half of pregnancy. The precise cause of pre-eclampsia is unknown, but it is characterized by disordered trophoblast invasion1 and abnormal placental angiogenesis2 In addition, an imbalance between prostacyclin, a vasodilator, and thromboxane, a vasoconstrictor, has been implicated in the pathogenesis of pre-eclampsia.3, 4 Although the mechanism is unknown, epidemiologic data suggests that populations with a high dietary intake of calcium have a decreased risk of pre-eclampsia.5,6 Based on this knowledge, a number of small and large clinical trials have been performed to evaluate the efficacy of antiplatelet agents and calcium for the prevention of pre-eclampsia. For each of the preventive strategies, the studies vary in size, inclusion/exclusion criteria, and primary outcomes.

With regards to aspirin, the smallest of the studies included 20 patients7, while the largest included 9364 patients.8 Among the calcium studies, the smallest enrolled 30 patients9 and the largest study enrolled over 8300 patients.10 For both aspirin and calcium, more of the smaller studies demonstrated a benefit while many of the larger ones did not. We make this point because these differences in the outcomes of published studies bring up the important topic of publication bias, which we will discuss in depth later in the article.

There were two large, randomized clinical trials performed in the United States by the NICHD-funded Maternal-Fetal Medicine Units Network to evaluate aspirin for the prevention of pre-eclampsia. The first tested the hypothesis that daily low dose aspirin could reduce the incidence of pre-eclampsia in healthy, nulliparous subjects.11 Over 3000 women were randomized to aspirin or placebo and this study revealed a relative risk of 0.70 (95% CI 0.60-1.0) for the development of pre-eclampsia, no difference in neonatal morbidity, and a slight increase in the incidence of placental abruption. The second study assessed low dose aspirin in 2539 women at high risk for pre-eclampsia12 and demonstrated an overall relative risk of 0.90 (95% CI 0.8-1.1) with no clear benefit in any subgroup and no evidence of adverse outcomes, including placental abruption. The largest study to date was conducted by the Collaborative Low-dose Aspirin Study in Pregnancy (CLASP) Group8 which randomized 9364 women who were felt to have a significant risk of pre-eclampsia to either low-dose aspirin or placebo for the prevention or treatment of pre-eclampsia and intrauterine growth restriction. Overall, 6.7 percent of women randomized to aspirin developed pre-eclampsia compared to 7.6 percent of women who received placebo, a non-significant 12 percent reduction in risk. Two other large trials also failed to demonstrate a benefit for aspirin in the prevention of pre-eclampsia.13,14

The PARIS collaborative group performed a patient level meta-analysis to assess whether there is a role for anti-platelet agents (mainly low dose aspirin) in the prevention of pre-eclampsia.15 This meta-analysis included individual patient level data from 32,217 women who were enrolled in 31 RCT's of antiplatelet agents for the prevention of pre-eclampsia. This meta-analysis generated relative risks of 0.90 (95% CI 0.84-0.97), for developing pre-eclampsia, 0.90 (95% CI 0.83-0.98) for delivery before 34 weeks' gestation and 0.90 (95%

CI 0.85-0.96) for a composite outcome of pre-eclampsia, delivery before 34 weeks, a small for gestational baby, stillbirth or death of the baby before discharge, or maternal death. The investigators found no difference in adverse outcomes including antepartum hemorrhage, post-partum hemorrhage, and placental abruption. Interestingly, no particular subgroup could be identified in which aspirin appeared to provide greater benefit. Based on a 10 percent reduction in the risk of pre-eclampsia the authors of this meta-analysis estimated the number of patients that a physician would need to treat to prevent one case of pre-eclampsia. The number needed to treat (NNT) takes into account how effective the treatment is (in this case a 10 percent reduction in risk) and also the baseline incidence of the disease. Given the 10 percent reduction in risk found in the PARIS study, if one assumes a baseline risk of pre-eclampsia of 18 percent (which would represent a very high risk population), 56 women would need to receive aspirin therapy to prevent one case of pre-eclampsia. If one assumes a baseline risk for pre-eclampsia of 6%, 167 women would need to be treated to prevent one case of pre-eclampsia. These results are similar to a traditional meta-analysis of antiplatelet agents for the prevention of pre-eclampsia performed by the Cochrane Database which identified a relative risk of 0.83 (95% CI 0.77-0.89) for the prevention of pre-eclampsia.16

Calcium supplementation for the prevention of pre-eclampsia has also been evaluated in a number of large and small trials, both in women with adequate and suboptimal calcium intake. The largest trial in women with adequate dietary intake of calcium was performed by the Calcium for the Prevention of Pre-Eclampsia (CPEP) group.17 This group assigned 4500 women with adequate dietary calcium intake to either calcium or placebo and found an overall relative risk for pre-eclampsia in the treatment group of 0.94 (95% CI 0.76-1.16). The World Health Organization (WHO) subsequently assessed calcium supplementation in 8325 women with inadequate dietary intake and found a RR for the development of pre-eclampsia of 0.91 (95% CI 0.69-1.19).18 Several secondary analyses of this trial were performed which revealed a RR of 0.76 (95% CI 0.66-0.89) for the development of a composite outcome of severe complications of pre-eclampsia, RR 0.8 (95% CI 0.7-0.91) for maternal morbidity/mortality, and RR 0.7 (95% CI 0.56-0.88) for neonatal mortality. The Cochrane Collaboration then combined 12 studies involving 15,000 women and found a RR of 0.70 (95% CI 0.57-0.86) for the development of hypertension in pregnancy and a RR of 0.48 (95% CI 0.33 to 0.69) for the development of pre-eclampsia. When they examined trials including patients with adequate dietary intake, there was no overall benefit from calcium supplementation (RR 0.62, 95% CI 0.32-1.20) whereas when data from 7 trials involving 10,154 women with inadequate dietary calcium intake was combined, the risk of developing pre-eclampsia was reduced with calcium supplementation (RR 0.36, 95% CI 0.18-0.70).19 We note that evidence of benefit in patients with low dietary calcium intake was centered in trials enrolling less than 300 patients.

Benefits and limitations of randomized controlled trials

Large, well done randomized controlled trials (RCT) are the gold standard for studies that change our clinical practice. The goal of any RCT is to design a study that is "big enough" so that a clinically significant effect is statistically significant but not "too big" so that clinically unimportant results become statistically significant. In addition, excessively large trials may be prohibitively expensive and the sample size needed may be large enough to

prohibit completing the study in a reasonable time frame. The sample size for any study is calculated based on the amount of Type I and Type II error the investigator is willing to tolerate, the expected number of outcomes in the placebo group, and the reduction in the primary outcome that the author wishes to be able to detect.

One strength of the RCT is that random allocation largely avoids the issue of confounding. However, there are important drawbacks of single large randomized clinical trials. First, some RCT's are so rigorous in their treatment and follow up that results may be "unique" to this setting. In addition, patients enrolled in clinical trials may have a baseline risk for the outcome of interest much greater than the average patient population, or may originate from very different populations. For example, some of the studies of calcium for the prevention of pre-eclampsia (while well-powered and methodologically sound) were performed in countries with very different diets than the US. A reasonable person may question whether these results can be applied to practice in the US and this reminds us that there are always concerns about the generalizability of large clinical trials. Another limitation is subtle variation across centers in subject enrollment into multicenter trials. One center may recruit lower-risk women enrolled from clinics where they receive routine care, while others using the same criteria may recruit higher-risk women referred for initial care. In addition, some trials may recruit women from different geographic areas, or even from different countries. These variations can be accounted for by strict inclusion/exclusion criteria and by stratifying randomization by center, but the reader of such articles should be aware of the potential impact of these considerations on trial outcome.

Oftentimes clinical trials are powered to detect a "composite outcome" rather than a single primary outcome. An example of this relates to the use of composite outcomes in prematurity research, which often combine severe RDS, NEC, IVH, and death into a single variable. Composite outcomes can be very helpful in that they reduce the estimated sample size in a clinical trial and make it easier to achieve statistically and clinically significant results. However, the use of a composite outcome limits the ability of any one study to detect differences in each of these important clinical outcomes individually. In addition, single large clinical trials may not be adequately powered to assess the frequency of rare but potentially important adverse events One example is the previously discussed NICHD-Maternal Fetal Medicine Units Network trial utilizing aspirin for the prevention of pre-eclampsia in nulliparous women.[11] This trial indicated that there was an increase in the risk of placental abruption in patients receiving low-dose aspirin versus placebo. Other randomized, controlled trials and large meta-analyses found rates of abruption in the placebo group which were similar to the rate in the aspirin group in the Network trial, suggesting that the initial association between aspirin and placental abruption was likely a chance association.

Also, single large clinical trials oftentimes may not have sufficient numbers of subjects in important subgroups. This is illustrated by the trials evaluating aspirin for the prevention of pre-eclampsia. Despite the large numbers of women enrolled in many of the studies there is still limited information available about important subgroups such as women with pre-existing renal disease.

One final concern with RCTs is the issue of publication bias that we alluded to earlier in the article. Publication bias occurs when the results of studies which report positive results are more likely to be published than studies with negative results. Studies which report significant results may result in a greater number of publications and may also be published in higher impact journals.[20] Although this type of bias might occur more frequently in observational studies than in randomized controlled trials[20] we posit that a small, positive randomized controlled trial would be easier to publish than a small, negative randomized controlled trial. As we will describe in the next section, publication bias can have an important impact on the outcomes of meta-analyses.

Benefits and limitations of meta-analysis

We have reviewed above the findings of a patient- level meta-analysis and traditional meta-analysis for therapies to prevent pre-eclampsia. We will now provide some of the background methodology for each type of study. Meta-analysis is a type of analysis in which results of a number of randomized clinical trials are statistically combined, and a "summary" measure of effect for a given treatment is ascertained. Many methods for meta-analysis are available, but the most commonly applied focus on the combination of published summary statistics, usually in some form of weighted average. This is the type of meta-analysis commonly performed by the Cochrane Collaboration. One important component of traditional meta-analysis is the process by which each study is assigned a "weight". A summary statistic is generated for each individual study to describe the observed intervention effect and a weight is then assigned depending on the precision of the effect size for that study.[21] Instead of simply assigning weight based on study size, the goal of this approach is to give more weight to the more precise studies. This is one way in which a small study with a high incidence of the outcome in question may contribute significantly to the overall results of meta-analysis.

In contrast, patient level meta-analysis refers to a process where individuals conducting a meta-analysis obtain individual data on each patient entered in all trials for central collection, processing, and analysis. This allows standard analysis to be performed and an overall result, based on the totality of the available evidence, to be calculated.[22] Reanalysis of all of the individual patients' data is widely considered to be the gold standard[23] and it has several advantages: it avoids biases associated with the use of summary statistics from separate studies, it allows the examination of data in detail, it eliminates dependence on already performed statistical analyses and instead allows information from each of the studies to be analyzed together, and time to event analyses can be conducted. In addition, adjustment for confounders and the search for differences in subgroups of patients based on characteristics of individual women that go beyond the summary results presented in the published trials can be performed. In some ways, patient level meta-analysis allows the authors to reconstruct the equivalent of a "mega-trial". This may prove particularly helpful in cases where the available randomized controlled trials are small and underpowered for the outcome of interest. However, even a patient level meta-analysis cannot address the problems associated with publication bias, differences in patient groups, and subtle differences in trial protocols and execution. One danger when combining multiple

adequately powered RCTs is that the large sample sizes associated with such a "mega-trial" may result in statistically significant results which may not be clinically significant.

One argument against meta-analysis is that it can combine patient populations that are dissimilar in one or more ways. In the case of aspirin for the prevention of pre-eclampsia studies differ in patient populations, definitions of pre-eclampsia, the type and dose of antiplatelet agent, and the timing of initiation of therapy. Some trials evaluating calcium for the prevention of pre-eclampsia enrolled women with low calcium intake; others enrolled women with normal dietary calcium intake. These variables must be considered when deciding whether to combine studies or to pool individual patient-level data and it is easy to see how such differences can markedly influence the results of a meta-analysis.

A second, somewhat related issue about meta-analysis is management of divergent results of studies. For example, consider two studies of equal size to evaluate a drug for preeclampsia where Study A found an improved outcome (RR= 0.5), while Study B reported worse outcome with treatment (RR=2.0). A meta-analysis of these studies would likely result in the misleading conclusion that treatment had no effect, either positive or negative, on outcome. (Figure 1). The epidemiological term for this is inter-study heterogeneity. Heterogeneity may not always be this obvious but it can be detected using a variety of statistical techniques. When heterogeneity is significant, we need to consider why heterogeneity is present, and whether or not combining the studies under consideration is appropriate. Of note, the Cochrane review assessing calcium for the prevention of pre-eclampsia detected significant heterogeneity among the studies. In this case heterogeneity seems to arise from smaller studies demonstrating greater benefit. This may have resulted from publication bias or it may result from the fact that smaller studies tend to enroll the highest risk patients.

A third point regarding meta-analysis relates to the quality of data going in- the so called "garbage in, garbage out" theory. This is an important concern that can have a significant impact on the results of meta-analyses. Schulz, et al analyzed a database of 250 obstetric trials from 33 meta-analyses and provided evidence that RCTs with inadequate blinding of patients and providers overestimated the intervention effect by 30 to 40 percent when compared to trials with adequate blinding.[24] The authors of a meta-analysis can address issues of study quality by carefully selecting and reporting inclusion and exclusion criteria as well as characteristics of included and excluded studies.[25]

For the examples of both aspirin and calcium for the prevention of pre-eclampsia we observe a trend that the earlier, smaller trials tend to provide more evidence of benefit that later, larger trials. This brings us to the point that meta-analysis has a crucial role in generating hypotheses for future RCTs, especially when only small trials exist. Importantly, one issue relating to meta-analysis of small trials was addressed in a study by LeLorier, et al. These authors identified 12 randomized, controlled trials enrolling 1000 patients or more. They were then able to identify 19 meta-analyses published on the same topic **prior** to publication of the large RCT. Outcomes of the 12 large RCTs were not predicted accurately by the prior meta-analyses 35 percent of the time.[26] This highlights the fact that meta-analysis cannot take the place of an adequately powered RCT, and that large clinical trials such as those

conducted for both aspirin and calcium for the prevention of pre-eclampsia are essential to clinical practice.

How do we make choices in clinical practice?

There are benefits to both meta-analysis and large randomized clinical trials, and each has its place in guiding clinical practice. In our opinion, when available, a well-done, adequately powered randomized clinical trial that enrolled a patient population comparable to the patient for whom a clinician is considering an intervention should be used to guide clinical management for several reasons.

First, we return to the example of calcium for the prevention of pre-eclampsia to make the point that there are many factors that can contribute to the overall results of a meta-analysis including the types of trials included and how much weight is assigned to each trial. The overall summary result of the calcium for the prevention of pre-eclampsia meta-analysis suggests that calcium supplementation is associated with a 52 percent reduction in the risk of pre-eclampsia. On closer examination, we observe that this result was centered on trials of women with inadequate dietary calcium intake and that the meta-analysis results for women with adequate dietary calcium intake agreed with the results of the CPEP trial which failed to demonstrate a benefit. This highlights the importance of knowing how summary results were obtained prior to adopting changes to clinical practice based on meta-analysis results. When we compare the results of the WHO trial, which failed to demonstrate a reduction in the risk of pre-eclampsia in women with inadequate dietary calcium intake to the meta-analysis, which detected a 67 percent reduction in the risk for pre-eclampsia in these women, we note that evidence of benefit was centered in those trials enrolling less than 260 women. While smaller trials may be well-conducted and methodologically sound, we continue to have concerns that significant publication bias in the case of smaller trials may influence the overall results of meta-analysis. We believe that meta-analysis has an important role in clinical situations where only small, underpowered clinical trials exist and conducting a large, more definitive trial is not feasible; in the examination of rare outcomes and hazards of therapy; and in identifying variation among studies. When possible, the hypotheses generated by meta-analysis of small, underpowered clinical should lay the groundwork to generate hypotheses for future large clinical trials.

In the case of aspirin for the primary prevention of pre-eclampsia there are a number of adequately powered clinical trials, encompassing a large number of patients in a wide variety of clinical settings that failed to demonstrate a benefit. The number of patients analyzed by the PARIS collaboration is impressive. However, we are concerned that such large numbers may result in conclusions that, although statistically significant, may represent clinically unimportant results. Of note, the analysis conducted by the PARIS collaboration also failed to identify a subgroup that clearly demonstrated a benefit from preventive therapy with aspirin and only a very mild improvement in outcomes. Even assuming a 10% reduction in the incidence of pre-eclampsia, in a low-risk patient population, a very large number needed to treat is present. If one considers a high risk population, the number needed to treat appears much lower based on the results of the meta-analysis. However, if we contrast this with the results of the large RCTs enrolling high risk

women that failed to demonstrate a benefit, we are inclined to use the results of the RCT to guide clinical management. While the results of any given RCT may not always be generalizable to all patient populations, the results of a meta-analysis cannot be preferentially applied to a clinical scenario where an adequately powered RCT exists.

Acknowledgments

References

1. Redman CW, Sargent IL. Latest advances in understanding pre-eclampsia. Science. 2005; 308:1592. [PubMed: 15947178]

2. Levine RJ, Maynard SE, Qian C, et al. Circulating angiogenic factors and the risk of pre-eclampsia. N Engl J Med. 2004; 350:672. [PubMed: 14764923]

3. Bussolino F, Benedetto C, Massobrio M, Camussi G. Maternal Vascular prostacyclin activity in pre-eclampsia. Lancet. 1980; 2:702. [PubMed: 6106815]

4. Masotti G, Galanti G, Poggesi L, et al. Differential inhibition of prostacyclin production and platelet aggregation by aspirin. Lancet. 1979; 2:1213. [PubMed: 92623]

5. Belizan JM, Villar J. The relationship between calcium intake and edema, proteinuria, and hypertension-gestosis: a hypothesis. Am J Clin Nutr. 1980; 33:2202–10. [PubMed: 6999886]

6. Belizan JM, Villar J, Repke J. The relationship between calcium intake and pregnancy-induced hypertension: up-to-date evidence. Am J Obstet Gynecol. 1988; 1598:898–902.

7. Kincaid-Smith P, North RA, Fairly KF, et al. Prevention of pre-eclampsia in women with renal disease: a prospective randomized trial of heparin and dipyridamole. Nephrology. 1995; 1:297.

8. CLASP (Collaborative Low-dose Aspirin Study in Pregnancy) Collaborative Group. CLASP: a randomized trial of low-dose aspirin for the prevention and treatment of pre-eclampsia among 9364 pregnant women. Lancet. 1994; 343:619. [PubMed: 7906809]

9. Niromanesh S, Laghii S, Mosavi-Jarrahi A. Supplementary calcium in prevention of pre-eclampsia. Int J Obstet Gynecol. 2001; 74:17–21.

10. Villar J, Abdel-Aleem H, Merialdi M, et al. World Health Organization randomized trial of calcium supplementation among low calcium intake pregnant women. Am J Obstet Gynecol. 2006; 194:639–49. [PubMed: 16522392]

11. Sibai BM, Caritis SN, Thom E, et al. Prevention of pre-eclampsia with low-dose aspirin in healthy, nulliparous pregnant women. N Engl J Med. 1993; 329:1213. [PubMed: 8413387]

12. Caritis S, Sibai B, Hauth J, et al. Low-dose aspirin to prevent pre-eclampsia in women at high risk. N Engl J Med. 1998; 338:701. [PubMed: 9494145]

13. Golding J, et al. A randomized trial of low dose aspirin for primiparae in pregnancy. BJOG. 1998; 105:293–299.

14. Rotchell YE, Cruickshank JK, Phillips Gay M, et al. Barbados low dose aspiring study in pregnancy (BLASP): a randomized trial for the prevention of pre-eclampsia and its complications. BJOG. 1998; 105:286.

15. PARIS Collaborative Group. Antiplatelet agents for prevention of pre-eclampsia: a meta-analysis of individual patient data. Lancet. 2007; 369:1791. [PubMed: 17512048]

16. Duley L, Henderson-Smart DJ, Meher S, King JF. Antiplatelet agents for preventing pre-eclampsia and its complications. Cochrane Database of Systematic Reviews. 2007; (2) Article No. CD004659. DOI:10.1002/14541858.

17. Levine RJ, Hauth JC, Curet LB, et al. Trial of calcium to prevent pre-eclampsia. N Engl J Med. 1997; 337:69–76. [PubMed: 9211675]

18. Villar J, Abdel-Aleem H, Merialdi M, et al. World Health Organization randomized trial of calcium supplementation among low calcium intake pregnant women. Am J Obstet Gynecol. 2006; 194:639–49. [PubMed: 16522392]

19. Hofmeyer GJ, Atallah AN, Duley L. Calcium Supplementation during pregnancy for preventing hypertensive disorders and related problems. Cochrane Database of Systematic Reviews. 2006; (3) Art No: CD001059.

20. Easterbrook PS, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. Lancet. 1991; 337:867–72. [PubMed: 1672966]

21. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions version 5.0.0. The Cochrane Collaboration. 2008updated February 2008Available from www.cochrane-handbook.org

22. Stewart LA, Clard MJ, Cochrane working group on meta-analysis using individual patient data. Practical methodology of meta-analysis using updated individual patient data. Statistics in Medicine. 1995; 14:2057. [PubMed: 8552887]

23. Simmonds MC, Higgins JP, Stewart LA, et al. Meta-analysis of individual patient data from randomized trials: a review of methods used in practice. Clinical Trials. 2005; 2:209. [PubMed: 16279144]

24. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical estimates of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA. 1995; 273:408–12. [PubMed: 7823387]

25. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. Lancet. 1999; 354:1896–1900. [PubMed: 10584742]

26. LeLorier J, Gregoire G, Benhaddad A, et al. Discrepancies between meta-analysis and subsequent large randomized, controlled trials. NEJM. 1997; 337:536–42. [PubMed: 9262498]
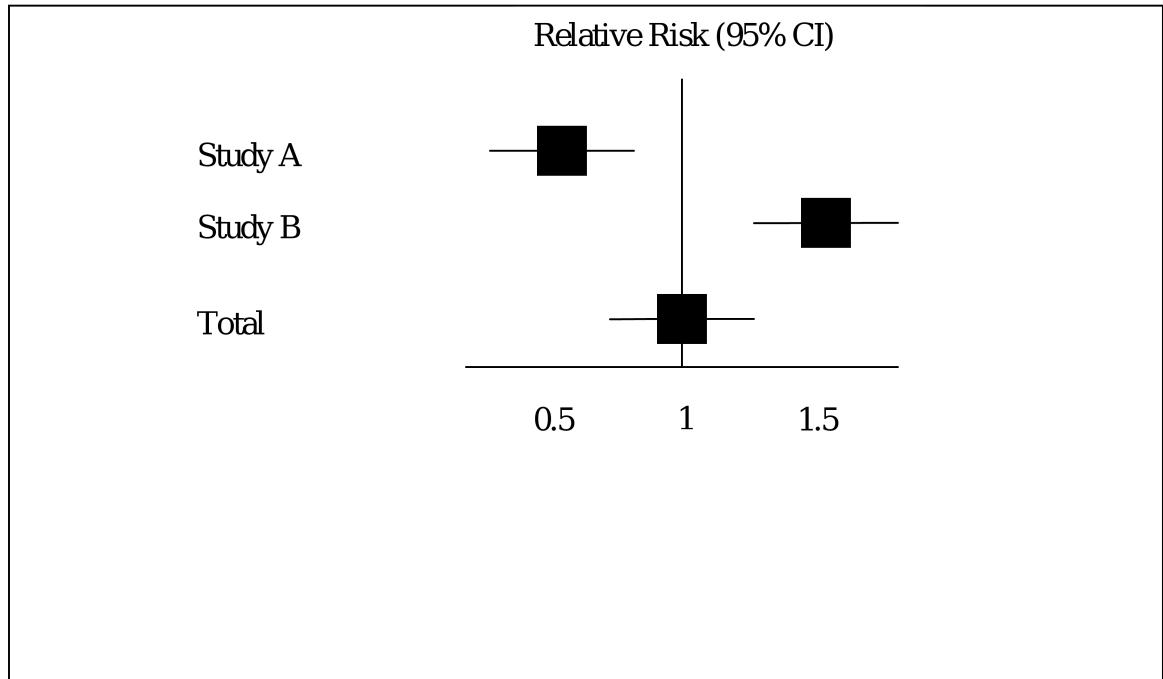
Figure 1.
Hypothetical example of inter-study heterogeneity and the effect on summary statistics